

DOCUMENT RESUME

ED 227 166

TM 830 188

AUTHOR Chambers, Barbara
TITLE Quality Control Review of Teacher-Made Tests.
INSTITUTION Cleveland Public Schools, Ohio. Dept. of Research, Development and Evaluation.
PUB DATE Jan 82
NOTE 30p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, Montreal, Quebec, Canada, April 11-15, 1983).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Elementary Secondary Education; English; French; Language Arts; Legislation; Mathematics; *Program Implementation; *Quality Control; Social Studies; *Teacher Made Tests; Test Construction; Test Reviews
IDENTIFIERS *Cleveland Public Schools OH

ABSTRACT

A process of reviewing teacher-made tests was developed to implement the court-approved subplan for a "procedure for continual quality control checks on teacher-constructed tests." The implementation process includes six steps: (1) distribution to teachers of Resource Materials Excerpted from Classroom Test Skills - a package to increase teacher skills in preparation, use, and analysis of teacher-made tests; (2) submission of a teacher-made test (chapter or unit test) to school files each semester; (3) collection of test files to forward to the Department of Research, Development and Evaluation (DRDE); (4) DRDE selection of tests from submitted files with each school represented; and (5) committee review and rating of tests by comparing each test to criteria based on the Classroom Test Skills guidelines; and (6) summarization and distribution of the test ratings to the individual schools. The reports of teacher-made test reviews are presented for the following six subject areas: upper elementary mathematics, high school English, primary language arts, 10th grade French, junior high social studies, and senior high mathematics. (Author/PN)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED227166

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as
received from the person or organization
originating it.
Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CLEVELAND PUBLIC SCHOOLS

Department of Research, Development and Evaluation

QUALITY CONTROL REVIEW OF TEACHER-MADE TESTS

Submitted to:

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

School Evaluation and Program Development Division

Entered in:

Category 7

Report on an Advance in the
Methodology of Policy Analysis,
Research, Evaluation,
Testing or Statistics.

Submitted by:

Barbara Chambers
Research Associate

January, 1982

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

B. Chambers

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

TM 830 188

DIVISION H

American Educational Research Association



ANNUAL EVALUATION REPORT AWARDS COMPETITION

This is to recognize

Barbara Chambers

for excellence in the field of educational evaluation reporting in the category

**Best Report On An Advance in the Methodology of Policy
Analysis, Research, Evaluation, Testing or Statistics**

for the entry

Quality Control Review of Teacher-Made Tests

In witness thereof this certificate is signed this 23rd *day of* March *, 1982.*

Shela M. Holly
Vice President, Division H

Barbara S. Prentice
Secretary, Division H

Carol Robinson
Awards Chairperson, Division H

Quality Control Review of Teacher-Made Tests

The process of reviewing teacher-made tests has been developed to implement the court approved subplan for a "procedure for continual quality control checks on teacher-constructed tests." This subplan addressed the component of the Remedial Order which stated, "... therefore, defendants are cautioned that they must insure that all tests, whether standardized, criterion referenced, or teacher-made, are developed, administered, and scored in a nondiscriminatory manner."

The implementation of the process includes the following procedures:

1. All teachers have received copies of the locally developed booklet, Resource Materials Excerpted From CLASSROOM TEST SKILLS.^{*} In addition, all teachers were requested to view thru the CLASSROOM TEST SKILLS slide/tape programs made available on educational TV.
2. Each semester every teacher submits one major (chapter or unit) teacher-made test to a file maintained by the principal in elementary schools or department chairpersons in secondary schools.
3. Principals (elementary) or department chairpersons (secondary) submit the most recent grade level/subject test file to subject supervisors upon request. The collected test files are then forwarded to the Department of Research, Development and Evaluation. (Test files from six grade levels or specific subjects are generally reviewed during each school year.)
4. The Department of Research, Development and Evaluation selects a random sample of tests from the submitted test files. At least one test is selected from each school. Tests are not identified by teacher or school, but are given a code number.
5. A committee review the sample of tests by comparing each randomly selected test to a set of criteria based on the CLASSROOM TEST SKILLS guidelines. Each test is rated by the use of the Teacher-Made Rating Form adapted from these guidelines. Membership on the committee generally includes the supervisor, two person from the Department of Research, Development and Evaluation, and two or more classroom teachers or department chairpersons.

^{*}CLASSROOM TEST SKILLS is a teacher-training package, locally developed through Title IV-C funding, to increase the skills of teachers in the preparation, use and analysis of teacher-made tests.

6. The ratings of the tests reviewed by the committee are are compiled and a report summarizing the ratings of all tests is prepared by the Department of Research, Development and Evaluation. This summary report, the rating sheet for the randomly selected test and the entire test file are forwarded to the supervisor for return to the individual schools.

This process of teacher-made test review has been completed for six subject areas to date. Copies of these Summary Reports follow:

- . Upper Elementary Mathematics
- . High School English
- . Primary Language Arts
- . Tenth Grade French
- . Junior High Social Studies
- . Senior High Mathematics

SUMMARY REPORT OF
REVIEW OF UPPER ELEMENTARY
TEACHER-MADE MATHEMATICS TESTS

Prepared by
Barbara Chambers
March 1981

Department of Research, Development and Evaluation
Cleveland Public Schools

SUMMARY REPORT of
REVIEW OF UPPER ELEMENTARY
TEACHER-MADE MATHEMATICS TESTS

TASK

The process of reviewing teacher-made tests has been developed to implement the court approved subplan for a "procedure for continual quality control checks on teacher-constructed tests." This subplan addressed the component of the Remedial Order which stated, "... therefore, defendants are cautioned that they must insure that all tests, whether standardized, criterion referenced, or teacher-made, are developed, administered, and scored in a nondiscriminatory manner."

PROCEDURE

The procedure for reviewing upper elementary teacher-made mathematics tests included the following:

1. Each teacher submitted one major test to a test file maintained in each school by the principal. Subject area of tests was at discretion of teachers.
2. Principals drew all upper elementary mathematics tests from the test file and forwarded them to the Supervisor of Elementary Mathematics. The tests were then forwarded to the Department of Research, Development and Evaluation.
3. A random sample of 45 tests was drawn from the tests submitted for review. At least one test was selected from each upper elementary school that had mathematics tests in their test files.
4. A committee reviewed the random sample of tests by comparing each test to a set of criteria based on the Classroom Test Skills guidelines. Comprising the review committee were: the Supervisor of Elementary Mathematics Department, an Assistant Supervisor, Research Associate, and Consultant Teacher from the Department of Research, Development and Evaluation; and two elementary mathematics teachers.
5. The upper elementary mathematics tests from the test file, the rating sheet for the randomly selected test, and this summary report were returned to each school.

SUMMARY OF FINDINGS:

Resulting from the review of 45 randomly selected upper elementary school teacher-made mathematics tests were the following findings:

- . The total number of items on the 45 tests consisted of 4.4% multiple choice items; 1.7% true-false; 5% matching; 89.4% problems or short-answer responses; and 0% essay questions.
- . All items were grouped by item type. That is, all multiple-choice items were grouped together, all problems, and so on.
- . Items were numbered consecutively throughout the majority of tests, however, approximately one-third of the tests numbered only rows of problems.
- . Most tests maintained the same item format within each test and every test kept all parts of a single item on one page.
- . With the exception of seven of the tests, directions were included for each item type. Frequently, however, the directions gave only operational directions, such as, add, subtract, or divide.
- . Thirty-five of the tests did not state any point values for test items.
- . The majority of tests were clearly written, neat in appearance and free of spelling, grammatical or apparent numerical error.
- . All tests included language and attitudes which were judged to be universal and non-discriminatory.
- . Multiple-choice items included on the tests generally stated the central problems in the stem, did not repeat phrases or terms, and included plausible distractors which were grammatically parallel and systematically listed.
- . True-false items included on the tests used clear, simple language, avoided specific determiners, were clearly true or false and avoided trick questions.
- . Matching exercises included on the tests explained the basis of matching and contained homogeneous brief lists. Several of the exercises, however, did not include extra distractors in matching lists.
- . Mathematical-type problems comprised the majority of items on the tests reviewed. The problems generally were clear and concise, directing the student to a definite brief answer. One or two of the problems, however, were impossible to solve based on the directions given students.
- . Short-answer response type questions usually employed the direct question approach, however, approximately 25% of those questions which included a "fill-in-the-blank" approach had response blanks of unequal length.

RECOMMENDATION

This report and the rated tests returned to each school should be reviewed by the principals with teachers in the school. The rating sheet and comments made regarding the reviewed tests should be viewed by teachers as an effort to help improve the teacher-made testing process in the schools. Because only one test was generally reviewed from each school's file, teachers must realize that this test may not have been representative of the quality of all teachers' tests within the school.

SUMMARY REPORT OF
REVIEW OF HIGH SCHOOL ENGLISH
TEACHER-MADE TESTS

Prepared by
Barbara Chambers
March 1981

Department of Research, Development and Evaluation
Cleveland Public Schools

SUMMARY REPORT of
REVIEW OF HIGH SCHOOL ENGLISH
TEACHER-MADE TESTS

TASK

The process of reviewing teacher-made tests has been developed to implement the court approved subplan for a "procedure for continual quality control checks on teacher-constructed tests." This subplan addressed the component of the Remedial Order which stated, "... therefore, defendants are cautioned that they must insure that all tests, whether standardized, criterion referenced, or teacher-made, are developed, administered, and scored in a nondiscriminatory manner."

PROCEDURE

The procedure for reviewing high-school English teacher-made tests included the following:

1. Each teacher submitted one major test to a test file maintained in each school by the department chairperson.
2. Department chairpersons forwarded tests from the test file to the Directing Supervisor of English who in turn forwarded them to the Department of Research, Development and Evaluation.
3. A random sample of 16 tests was drawn from the test files for review. At least one test was selected from each high school.
4. A committee reviewed the random sample of tests by comparing each test to a set of criteria based on the Classroom Test Skills guidelines. Comprising the review committee were: the Directing Supervisor of English, two Research Associates from the Department of Research, Development and Evaluation, a senior high school English Department Chairperson, and a senior high school English teacher.
5. The test file, the rating sheet for the randomly selected test, and this summary report were returned to each high school.

SUMMARY OF FINDINGS:

Resulting from the review of 16 randomly selected senior high school teacher-made English tests were the following findings:

- The total number of items on the 16 tests consisted of 21% multiple choice items; 16% true-false; 32% matching, 26% completion (short-answer response); and 5% essay questions.

- . Thirteen of the tests had items grouped by item-type. That is, all multiple-choice items were grouped together, all true-false, and so on.
- . Most tests were divided into parts and numbered within the parts rather than numbered consecutively throughout entire test.
- . Five of the tests did not include directions for each item-type, while 11 did not state consistent point values for each item-type.
- . The majority of tests were clearly written and neat in appearance, however, 7 tests contained spelling or grammatical errors.
- . All tests included language which was universal and non-discriminatory.
- . Two tests included attitudes which were judged to possibly be non-universal. One referred to "gang wars" and the other generalized beliefs of Puritans.
- . Multiple-choice items included on the tests generally stated the central problems in the stem, did not repeat phrases or terms and included plausible distractors. Most lists of distractors, however, were not grammatically parallel nor listed in some systematic order.
- . True-false items included on the tests used clear, simple language, avoided specific determiners, were clearly true or false and avoided trick questions.
- . Matching exercises included on the tests usually explained the basis for matching and contained homogeneous, brief lists. Several of the tests, however, did not indicate if response could be used more than once.
- . Completion (short-answer response) items included on the tests generally did not employ the direct question approach. Rather the questions included blanks of varying length, appearing in various positions within the question. These variations of item writing styles often resulted in ambiguous questions which may have had more than one correct answer.
- . Essay questions included on the test avoided terms such as "list," "how," or "how many" and generally indicated the point value of each question. Several of the tests reviewed, however, included general, all encompassing questions which did not direct the student to the desired response.

RECOMMENDATION

This report and the rated tests returned to each school should be reviewed by the English Department Chairperson with teachers in the department. The rating sheet and comments made regarding the reviewed test should

be viewed by teachers as an effort to help improve the teacher-made testing process in the schools. Because only one test was generally reviewed from each school's file, teachers must realize that this test may not have been representative of the quality of all teachers' tests within the department.

SUMMARY REPORT OF
REVIEW OF PRIMARY
TEACHER-MADE LANGUAGE ARTS TESTS

Prepared by
Barbara Chambers
April 1981

Department of Research, Development and Evaluation
Cleveland Public Schools

SUMMARY REPORT OF
REVIEW OF PRIMARY
TEACHER-MADE LANGUAGE ARTS TESTS

TASK

The process of reviewing teacher-made tests has been developed to implement the court approved subplan for a "procedure for continual quality control checks on teacher-constructed tests." This subplan addressed the component of the Remedial Order which stated, "... therefore, defendants are cautioned that they must insure that all tests, whether standardized, criterion referenced, or teacher-made, are developed, administered, and scored in a nondiscriminatory manner."

PROCEDURE

The procedure for reviewing primary teacher-made Language Arts tests included the following:

1. Each teacher submitted one major test to a test file maintained in each school by the principal. Subject area of tests was at discretion of teachers.
2. Principals drew all primary Language Arts tests from the test file and forwarded them to the Directing Supervisor of Language Arts. The tests were then forwarded to the Department of Research, Development and Evaluation.
3. A random sample of 42 tests was drawn from the tests submitted for review. At least one test was selected from each primary school that had Language Arts tests in their test files.
4. A seven-member committee reviewed the random sample of tests by comparing each test to a set of criteria based on the Classroom Test Skills guidelines. Comprising the Review Committee were: The Directing Supervisor and Supervisor of Language Arts; the Assistant Supervisor of the Listening Post Program; a Research Associate and Consultant Teacher from the Department of Research, Development and Evaluation; and two primary Language Arts teachers.
5. The primary Language Arts tests from the test file, the rating sheet for the randomly selected test, and this summary report were forwarded to the Directing Supervisor of Language Arts for return to the individual schools.

SUMMARY OF FINDINGS

Resulting from the review of 42 randomly selected primary level teacher-made Language Arts tests were the following findings:

Part I

- . All tests items on tests were grouped by item-type. That is, all sound or picture identification items, ordering items, prefix or suffix addition items, etc., were grouped together.
- . Approximately one-half of the tests numbered items consecutively throughout the entire test.
- . All items of a particular item-type were of consistent format on all the tests.
- . The majority of tests were single page documents, therefore, all parts of the exercises were kept on a single page.
- . Clear directions were included for all but six of the tests.
- . Only one of the tests indicated point values for test items.
- . The vast majority of the tests were typed or clearly handwritten, neat in appearance, and free of spelling or grammatical errors.
- . All tests included language which is universal and non-discriminatory.
- . All test, with the exception of one, reflected attitudes which are universal and non-discriminatory. The one exception included sentences geared only toward boys.

Part II

Generally, primary level teacher-made tests do not include items which can specifically be classified as multiple-choice, true-false, matching, short answer response (completion) or essay. Rather primary tests, which might better be called worksheets, tend to be reinforcement exercises which focus on sound identification, ordering and classifying processes, prefix and suffix addition. The primary Language Arts tests reviewed by the committee typified this expected item make-up. For this reason, Part II of the Rating Form was not applicable for use with the primary tests reviewed.

Part III

The General Comments made by the test raters are summarized below in terms of the nature and frequency of the comments:

- . Directions
 - 3 commending good directions
 - 15 noting incomplete or missing directions
- . Neatness and ease of reading
 - 6 commending the quality
 - 1 noting lack of quality
- . Appropriateness of test for grade level
 - 3 commending appropriateness
 - 2 noting inappropriateness
- . Size of type
 - 1 commending the size
 - 2 indicating type too small
- . Other
 - 9 saying "Good Test"
 - 7 noting errors in individual test items
 - 2 noting omission of a place for name and date
 - 2 indicating the need to number items
 - 1 indicating the need to include point values with items
 - 1 questioning choice of item format

RECOMMENDATIONS

This report and the rated tests returned to each school should be reviewed by the principal with teachers in the school. The rating sheet and comments regarding the reviewed test should be viewed by teachers as an effort to help improve the teacher-made testing process in the schools. Because only one test was generally reviewed from each school's file, teachers must realize that this test may not have been representative of the quality of all teachers' tests within the school.

SUMMARY REPORT OF
REVIEW OF 10th GRADE
TEACHER-MADE FRENCH TESTS

Prepared by
Barbara Chambers
May 1981

Department of Research, Development and Evaluation
Cleveland Public Schools

SUMMARY REPORT OF
REVIEW OF 10th GRADE
TEACHER-MADE FRENCH TESTS

TASK

The process of reviewing teacher-made tests has been developed to implement the court approved subplan for a "procedure for continual quality control checks on teacher-constructed tests." This subplan addressed the component of the Remedial Order which stated, "...therefore, defendants are cautioned that they must insure that all tests, whether standardized, criterion referenced, or teacher-made, are developed, administered, and scored in a nondiscriminatory manner."

PROCEDURE

The procedure for reviewing 10th grade French teacher-made tests included the following:

1. Each teacher submitted one major test to a test file maintained in each school by the department chairperson.
2. Department chairpersons forwarded tests from the test file to the Assistant Supervisor of Foreign Languages who in turn forwarded them to the Department of Research, Development and Evaluation.
3. A random sample of 12 tests was drawn from the test files for review. At least one test was selected from each high school that offers the course.
4. A committee reviewed the random sample of tests by comparing each test to a set of criteria based on the Classroom Test Skills guidelines. Comprising the review committee were: the Assistant Supervisor of Foreign Languages, two Research Associates from the Department of Research, Development and Evaluation, a senior high school Foreign Languages Chairperson, and a senior high school French teacher.
5. The test file, the rating sheet for the randomly selected test, and this summary report were forwarded to the Assistant Supervisor of Foreign Languages for return to individual schools.

SUMMARY OF FINDINGS:

Resulting from the review of 12 randomly selected 10th grade teacher-made French tests were the following findings:

PART I - Rate Entire Test

- . All test items were grouped by item-type. That is, all translation, matching, completion, essay or multiple choice items were grouped together.
- . Less than one-half of the test (5) numbered items consecutively throughout the entire test.
- . Items within each item-type retained the same format on 10 of the tests.
- . The majority of tests were single page documents, therefore, all parts of the exercises were kept on a single page.
- . Directions were included on 10 of the tests.
- . Only two of the tests include point valued for test items.
- . Nine tests were typed or clearly hand written neat in appearance.
- . All tests were free of spelling and grammatical errors.
- . All tests, with the exception of one, reflected attitudes which are universal and non-discriminatory. The one exception included questions which contained only male names.

PART II - Rate Item-Types

- . The 395 items on the 12 tests (average of 35 items per test) consisted of 74% classified as completion or short-answer response, 21% multiple-choice, 4% matching and 1% essay; no true-false items were included.
- . Multiple-choice items included on the tests had plausible distractors and listed responses in some systematic order. Because many of the items classified as multiple-choice were of the type directing students to select the correct form of a verb or translation of a word, it was not applicable to rate the quality of the stem.
- . Matching exercises included on the tests explained the basis for matching, contained homogeneous, brief lists and included extra distractors.

- . The majority of completion or short-answer responses included on the tests directed students to supply verb forms or translations. The items therefore did not employ the direct question approach, but did elicit brief, definite answers.
- . The items classified as essay type were questions directing students to write brief compositions in response to structured questions.

PART III - General Comments

The General Comments made by the test raters are summarized below in terms of the nature and frequency of the comments:

- . Directions
 - 1 commending good directions
 - 7 noting incomplete or missing directions
- . Length of test
 - 1 commending appropriateness
 - 3 saying too comprehensive
- . Point value of items
 - 1 commending inclusion of point value
 - 8 noting missing point values
- . Quality of test copy
 - 3 noting poor quality of copy
- . Other
 - 6 saying "Good Test"

RECOMMENDATIONS

This report and the rated tests returned to each school should be reviewed by the department chairperson with teachers in the school. The rating sheet and comments regarding the reviewed test should be viewed by teachers as an effort to help improve the teacher-made testing process in the schools. Because only one test was generally reviewed from each school's file, teachers must realize that this test may not have been representative of the quality of all teachers' tests within the department.

**SUMMARY REPORT OF
REVIEW OF JUNIOR HIGH
TEACHER-MADE SOCIAL STUDIES TESTS**

**Prepared by
Barbara Chambers
June 1981**

**Department of Research, Development and Evaluation
Cleveland Public Schools**

SUMMARY REPORT OF
REVIEW OF JUNIOR HIGH
TEACHER-MADE SOCIAL STUDIES TESTS

TASK

The process of reviewing teacher-made tests has been developed to implement the court approved subplan for a "procedure for continual quality control checks on teacher-constructed tests." This subplan addressed the component of the Remedial Order which stated, "...therefore, defendants are cautioned that they must insure that all tests, whether standardized, criterion referenced, or teacher-made, are developed, administered, and scored in a nondiscriminatory manner."

PROCEDURE

The procedure for reviewing junior high Social Studies teacher-made tests included the following:

1. Each teacher submitted one major test to a test file maintained in each school by the department chairperson
2. Department chairpersons forwarded tests from the test file to the Supervisor of Social Studies who in turn forwarded them to the Department of Research, Development and Evaluation.
3. A random sample of 23 tests was drawn from the test files for review. At least one test was selected from each high school that offers the course. (Note: All tests submitted from one school were declared to be commercially prepared and therefore a sample test was not rated.)
4. A committee reviewed the random sample of tests by comparing each test to a set of criteria based on the Classroom Test Skills guidelines. Comprising the review committee were: the Supervisor of Social Studies, two Research Associates from the Department of Research, Development and Evaluation, two junior high school Social Studies Chairperson, and a junior high school Social Studies teacher.
5. The test file, the rating sheet for the randomly selected test, and this summary report were forwarded to the Supervisor of Social Studies for return to individual schools.

SUMMARY OF FINDINGS:

Resulting from the review of 23 randomly selected junior high teacher-made Social Studies tests were the following findings:

Part I - RATE ENTIRE TEST

- . Items were grouped by item-type on all tests but one. That is, all multiple-choice, matching, essay, true-false, or completion items were grouped together.
- . Less than one-half of the tests (11) numbered items consecutively throughout the entire test.
- . Items within each item-type retained the same format on 18 of the tests.
- . Six of the tests did not keep all parts of a single item on the same page
- . Directions were included on 10 of the tests.
- . Only two of the tests included point values for test items.
- . The vast majority of tests (19) were typed or clearly handwritten, neat in appearance, and free of spelling and grammatical errors.
- . All tests included language and attitudes which are universal and non-discriminatory.

Part II - RATE ITEM-TYPE

- . The 798 items on the test rated (average of 35 items per test) consisted of 38% classified as matching, 22% completion or short-answer response, 19% true-false, 15% multiple-choice, 3% essay, and 3% other.
- . Multiple-choice items included on the tests had plausible distractors, contained grammatically parallel lists of responses, avoided repetition of phrases or terms, and stated the problem clearly in the stem. Responses were consistently listed in a systematic order on only one-third of the tests which contained multiple-choice items.
- . True-false items included on the tests used clear, simple language and avoided specific determiners and trick questions. Items were judged to be for the majority of items clearly true or false.
- . Of the 16 matching exercises included on the tests, seven explained the basis for matching and included homogeneous lists. Nine of the exercises maintained brief lists and included extra distractors.
- . The majority of completion or short-answer responses included on the tests elicited brief, definite answers and appeared to avoid verbatim textbook language. Five of the

16 tests containing completion or short-answer response items, employed the the direct question approach. Four of the tests consistently maintained response blanks of equal lengthh.

- . The 9 tests that included essay items avoided general, all encompassing questions and constructed the items simply and clearly to direct students to the desired response. Four of the tests, however, incorrectly classified items containing words such as "list" or "who" as essays. These items actually were short-answer response items.

Part III - GENERAL COMMENTS

The General Comments made by the test raters are summarized below in terms of the nature and frequency of the comments:

- . Directions
 - 1 commending good directions
 - 6 noting incomplete or missing directions
- . Format of Test
 - 3 commending the format
 - 5 questioning the format
- . Ease of Reading
 - 3 commending the quality
 - 1 noting difficulty
- . Item-Type
 - 3 noting incorrect classification of items
 - 3 suggesting an alternative item-type
 - 1 suggesting items within item-type be grouped by topic
 - 1 noting that the same topic of information was included in two types of items
 - 1 noting that items were not grouped by item-type
- . Construction of Items
 - 1 suggesting improvements in true-false items
 - 4 suggesting improvements in essay items
 - 7 suggesting improvements in matching exercises
- . Other
 - 4 suggesting consecutive numbering of all items
 - 4 suggesting that point values be included
 - 4 noting spelling or punctuation errors
 - 4 saying "Good Test"

RECOMMENDATIONS

This report and the rated tests returned to each school should be reviewed by the department chairperson with teachers in the school. The rating sheet and comments regarding the reviewed test should be viewed by teachers as an effort to help improve the teacher-made testing process in the schools. Because only one test was generally reviewed from each school's file, teachers must realize that this test may not have been representative of the quality of all teachers' tests within the department.

-- SUMMARY REPORT --
REVIEW OF SENIOR HIGH SCHOOL
TEACHER-MADE MATHEMATICS TESTS

Prepared by
Barbara Chambers
November 1981

Department of Research, Development and Evaluation
Cleveland Public Schools

SUMMARY REPORT
REVIEW OF SENIOR HIGH SCHOOL
TEACHER-MADE MATHEMATICS TESTS

TASK

The process of reviewing teacher-made tests has been developed to implement the court approved subplan for a "procedure for continual quality control checks on teacher-constructed tests." This subplan addressed the component of the Remedial Order which stated, "...therefore, defendants are cautioned that they must insure that all tests, whether standardized, criterion referenced, or teacher-made, are developed, administered, and scored in a nondiscriminatory manner."

PROCEDURE

The procedure for reviewing senior high school mathematics teacher-made tests included the following:

1. Each teacher submitted one major test to a test file maintained in each school by the department chairperson.
2. Department chairpersons forwarded tests from the test file to the Supervisor of Mathematics who in turn forwarded them to the Department of Research, Development and Evaluation.
3. A random sample of 20 tests was drawn from the test files for review. At least one test was selected from each high school.
4. A committee reviewed the random sample of tests by comparing each test to a set of criteria based on the Classroom Test Skills guidelines. Comprising the review committee were: the Supervisor of Mathematics, two Research Associates from the Department of Research, Development and Evaluation, and two senior high school Mathematics department chairpersons.
5. The test file, the rating sheet for the randomly selected test, and this summary report were forwarded to the Supervisor of Mathematics for return to individual schools.

SUMMARY OF FINDINGS:

Resulting from the review of 20 randomly selected senior high school teacher-made mathematics tests were the following findings:

PART I. RATING OF TEST DOCUMENT

- . Seventeen (80%) grouped items by item-type and numbered items consecutively throughout the entire test.
- . Sixteen (80%) were consistent in format within item types. Fifteen (75%) kept all parts of a single item on one page.
- . Fourteen (70%) had complete directions for each item type. One (5%) included consistent point values.
- . Eighteen (90%) were typed or clearly handwritten, neat in appearance. Seventeen (85%) were free of errors--spelling, grammatical or apparent numerical error, however, two tests numbered test items incorrectly.
- . All tests included language which was judged to be universal and non-discriminatory. Fifteen (75%) reflected attitudes which are universal and non-discriminatory; this trait could not be determined on four tests and was not applicable to one test. The term "man" was used in one test item.

PART II. RATING OF TEST ITEMS BY ITEM-TYPE

- . The 430 items on the 20 tests (average of 21.5 items per test) consisted of 17.2% classified as completion or short-answer response, 10.5% matching, .9% as proofs and 71.4% as numerical problems; no true-false, matching, or essay items were included.
- . Matching exercises (45) included on the tests explained the basis for matching, contained homogeneous, brief lists and included extra distractors.

PART III. GENERAL COMMENTS REGARDING THE QUALITY OF ENTIRE TEST

- . Three (15%) questioned the vast number of concepts included on a single test.
- . Five (25%) related to the need for more explicit directions.
- . Four (20%) gave suggestions for improving the format.
- . One (5%) regarded the test to be too long.
- . One (5%) suggested including point values for items.
- . Nine (45%) commended the general quality of the tests.

RECOMMENDED USE OF THIS REPORT

This report and the rated tests returned to each school should be reviewed by the department chairperson with teachers in the school. The rating sheet and comments regarding the reviewed test should be viewed by teachers as an effort to help improve the teacher-made testing process in the schools. Because only one test was generally reviewed from each school's file, teachers must realize that this test may not have been representative of the quality of all teachers' tests within the department.